

LOWER BOUNDS FOR POLYNOMIALS WITH SIMPLEX NEWTON POLYTOPES BASED ON GEOMETRIC PROGRAMMING

SADIK ILIMAN AND TIMO DE WOLFF

ABSTRACT. In this article, we propose a geometric programming method in order to compute lower bounds for real polynomials. We provide a new sufficient criterion on the coefficients of a polynomial to be nonnegative as well as on the coefficients and the support to have a sum of binomial squares representation. It generalizes all previous criteria by Lasserre, Ghasemi, Marshall, Fidalgo and Kovacec from scaled standard simplex Newton polytopes to arbitrary simplex Newton polytopes.

This generalization yields a geometric programming approach for computing lower bounds for polynomials that significantly extends the one proposed by Ghasemi and Marshall in [6]. Furthermore, it shows that geometric programming is strongly related to nonnegativity certificates given by sums of nonnegative circuit polynomials, which were recently introduced by the authors in [10].

1. INTRODUCTION

Finding lower bounds for real polynomials is a central problem in polynomial optimization. Several well known approaches to this problem work well in small dimension or with additional structure enforced on the polynomials. The best known lower bounds are provided by *Lasserre relaxations* using semidefinite programming. In spite of the fact that the optimal value of a semidefinite program can be computed in time polynomial up to an additive error, the size of such programs grows rapidly with the number of variables or degree of the polynomials. Therefore, recently, there is much interest in finding lower bounds for polynomials using the alternative approach of *geometric programming* (see (4.1) for a formal definition). In recent works [6, 7] several important facts are shown for polynomial optimization via geometric programming. Two key observations are the following ones:

- (1) For general polynomials lower bounds based on geometric programming are seemingly not as good as bounds obtained by semidefinite programming.
- (2) Even higher dimensional examples can often be solved quite fast with geometric programming whereas semidefinite programs do not yield an output at all due to the too high dimension resp. degree of polynomials.

A *global polynomial optimization problem* for some $f \in \mathbb{R}[x]_{2d} = \mathbb{R}[x_1, \dots, x_n]_{2d}$ of degree $2d$ is given by

2010 *Mathematics Subject Classification.* 12D15, 14P99, 52B20, 90C25.

Key words and phrases. Geometric programming, lower bound, nonnegative polynomials, simplex, sparsity, sums of squares.

$$f^* = \inf\{f(x) : x \in \mathbb{R}^n\} = \sup\{\lambda \in \mathbb{R} : f - \lambda \geq 0\}.$$

It is well known that computing f^* is NP-hard in general ([2]). Relaxing the nonnegativity condition to a sum of squares condition, a lower bound for f^* based on semidefinite programming is given by

$$f_{sos} = \sup \left\{ \lambda \in \mathbb{R} : f - \lambda = \sum_{i=1}^k q_i^2 \text{ for some } q_i \in \mathbb{R}[x] \right\}$$

and hence $f_{sos} \leq f^*$ ([12]). An open key problem in polynomial optimization is to analyze the gap $f^* - f_{sos}$. Very little is known about this gap beyond the cases where it always vanishes, which, by Hilbert's Theorem, happens for $(n, 2d) \in \{(1, 2d), (n, 2), (2, 4)\}$ ([9]). For an overview see about the topic see, e.g., [1, 12, 13].

In this article, we extend results in [6] by Ghasemi and Marshall in order to provide lower bounds for polynomials using *geometric programs*, which can be solved in time polynomial using interior point methods ([14]). We denote these lower bounds by f_{gp} .

In fact, this extension relies on the key observation that nonnegativity can not only be certified via sums of squares, but also via *sums of nonnegative circuit polynomials*, which were recently introduced by the authors in [10]. In [5], Fidalgo and Kovacec provide nonnegativity resp. sums of squares certificates for a class of polynomials, which have the scaled standard simplex $\text{conv}\{0, 2d \cdot e_1, \dots, 2d \cdot e_n\}$ as Newton polytopes. In [6] Ghasemi and Marshall show that these certificates can be translated into geometric programs. But since the certificates in [5] are special instances of the authors ones in [10], it is self-evident to ask, whether the translation into geometric programs can also be generalized. The purpose of this article is to show that this is indeed the case.

As main theoretical results we contribute in Section 3 some easily checkable criteria on the coefficients of a polynomial with simplex Newton polytope, which imply that the polynomial is a *sum of nonnegative circuit polynomials (SONC)*. Additionally, we provide some easily checkable criteria on the support of the polynomial, which imply that such an SONC additionally is a sum of binomial squares (see Theorems 3.1 and 3.2; see Section 2 for a formal definition of SONC).

The key observation is that, as in [6], these criteria can be translated into a geometric optimization problem (Corollary 4.2). As a surprising fact we show in Corollary 3.4 that for very rich classes of polynomials with simplex Newton polytope, the optimal value f_{gp} satisfies $f_{gp} \geq f_{sos}$ – in contrast to the general observation by Ghasemi in Marshall ([6, 7]), which we outlined in the beginning. Additionally, the computation of f_{gp} is much faster than in the corresponding semidefinite optimization problem.

Using the geometric programming software package GPPOSY for MATLAB, we demonstrate the capabilities of our results on the basis of different examples. A major observation is the fact that the bounds f_{gp} and f_{sos} are not comparable in general.

2. PRELIMINARIES

We consider polynomials $f \in \mathbb{R}[x]_{2d}$ of the form $f = \sum_{\alpha \in \mathbb{N}_{2d}^n} f_\alpha x^\alpha$ with $f_\alpha \in \mathbb{R}$, $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ and Newton polytope $\text{New}(f) = \text{conv}\{\alpha \in \mathbb{N}_{2d}^n : f_\alpha \neq 0\}$. Throughout this article we assume that $\text{New}(f)$ is a simplex with vertex set $\{0, \alpha(1), \dots, \alpha(n)\} \subset \mathbb{N}_{2d}^n$ and corresponding coefficients $f_0, f_{\alpha(j)} > 0$ for $1 \leq j \leq n$. Following existing literature [5, 6, 11], we define

$$\Omega(f) = \{\alpha \in \mathbb{N}_{2d}^n : f_\alpha \neq 0\} \setminus \{0, \alpha(1), \dots, \alpha(n)\}.$$

Hence, we have a decomposition

$$f = f_0 + \sum_{j=1}^n f_{\alpha(j)} x^{\alpha(j)} + \sum_{\alpha \in \Omega(f)} f_\alpha x^\alpha$$

where $f_0 > 0$ is the constant term in f . Let

$$\begin{aligned} \Delta(f) &= \{\alpha \in \Omega(f) : f_\alpha x^\alpha \text{ is not a square}\} \\ &= \{\alpha \in \Omega(f) : f_\alpha < 0 \text{ or } \alpha_i \text{ is odd for some } 1 \leq i \leq n\}. \end{aligned}$$

In order to describe the degree of our extension of the work [6] in more detail, we briefly review results from the authors in [10].

Theorem 2.1. *Let $f = \lambda_0 + \sum_{j=1}^n f_{\alpha(j)} x^{\alpha(j)} + c \cdot x^y \in \mathbb{R}[x]_{2d}$ with $f_{\alpha(j)} > 0$ be a polynomial such that $\text{New}(f) = \text{conv}\{0, \alpha(1), \dots, \alpha(n)\}$ is a simplex and $\text{supp}(f) = \{0, \alpha(1), \dots, \alpha(n), y\}$ with $y \in \text{int}(\text{New}(f)) \cap \mathbb{N}^n$. Then the following are equivalent.*

- (1) f is nonnegative.
- (2) $|c| \leq \Theta_f$ and $y \notin (2\mathbb{N})^n$ or $c \geq -\Theta_f$ and $y \in (2\mathbb{N})^n$

where $\Theta_f = \prod_{j=0}^n \left(\frac{f_{\alpha(j)}}{\lambda_j} \right)^{\lambda_j}$ with $\sum_{j=0}^n \lambda_j \alpha(j) = y$, $\lambda_j > 0$, $\sum_{j=0}^n \lambda_j = 1$.

We denote polynomials considered in the above theorem as *circuit polynomials*, since they are supported on a *circuit*, i.e., a minimally affine dependent set. As an immediate consequence, we have that if a real polynomial f is a *sum of nonnegative circuit polynomials (SONC)*, then this is a certificate for nonnegativity of f (see [10] for further details on the cone of SONC's).

Corollary 2.2. *Let $f \in \mathbb{R}[x]_{2d}$. Then f is nonnegative if there exist nonnegative circuit polynomials $g_1, \dots, g_s \in \mathbb{R}[x]$ (in the sense of Theorem 2.1) and $\mu_j \geq 0$ such that $f = \sum_{j=1}^s \mu_j g_j$.*

For general nonnegative polynomials, whether a SONC decomposition exists resp. how it can be computed is unclear so far. But for the case of polynomials f with simplex Newton polytopes it was shown in [10, Corollary 7.4.] that the existence of a SONC decomposition is equivalent to nonnegativity of f if there exists an orthant where all terms of f except those corresponding to vertices have a negative sign.

Theorem 2.3. *Let $f = f_0 + \sum_{j=1}^n f_{\alpha(j)} x^{\alpha(j)} + \sum_{\alpha \in \Omega(f)} f_\alpha x^\alpha$ be nonnegative with $f_{\alpha(j)} \in \mathbb{R}_{>0}$ and $f_\alpha \in \mathbb{R}^*$ such that $\text{New}(f) = \Delta = \text{conv}\{0, \alpha(1), \dots, \alpha(n)\}$ is a simplex and*

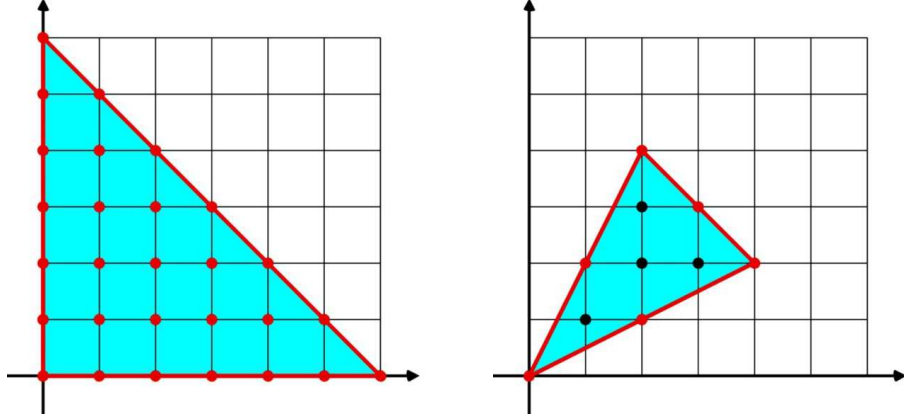


FIGURE 1. On the left: The H -simplex $\text{conv}\{(0,0), (6,0), (0,6)\} \subset \mathbb{R}^2$. On the right: The M -simplex $\text{conv}\{0, (2,4), (4,2)\} \subset \mathbb{R}^2$. The red (light) points are the lattice points contained in the corresponding sets Δ^* .

$\Omega(f) \subseteq (\text{int}(\Delta) \cap \mathbb{N}^n)$. If there exists a vector $v \in (\mathbb{R}^*)^n$ such that $f_\alpha v^\alpha < 0$ for all $\alpha \in \Omega(f)$, then f is SONC.

In order to talk about results concerning sums of squares, we need to define $\hat{\Delta}$ -mediated sets first introduced by Reznick in [15] (see also [10]).

Definition 2.4. Let $\hat{\Delta} = \{0, \alpha(1), \dots, \alpha(n)\} \subset (2\mathbb{N})^n$ be such that $\text{conv}(\hat{\Delta})$ is a simplex and let $L \subset \text{conv}(\hat{\Delta}) \cap \mathbb{Z}^n$.

- (1) Define $A(L) = \{\frac{1}{2}(s+t) \in \mathbb{Z}^n : s, t \in L\}$ and $\overline{A}(L) = \{\frac{1}{2}(s+t) \in \mathbb{Z}^n : s \neq t, s, t \in L\}$ as the set of averages of even resp. distinct even points in $\text{conv}(L) \cap \mathbb{Z}^n$.
- (2) We say that L is $\hat{\Delta}$ -mediated, if

$$\hat{\Delta} \subset L \subset \overline{A}(L) \cup \hat{\Delta},$$

i.e., every $\beta \in L \setminus \hat{\Delta}$ is an average of two distinct even points in L .

The key fact about $\hat{\Delta}$ -mediated sets is that there always exists a maximal one w.r.t. inclusion.

Theorem 2.5 (Reznick [15]). *There is a $\hat{\Delta}$ -mediated set Δ^* satisfying $A(\hat{\Delta}) \subseteq \Delta^* \subseteq (\text{conv}(\hat{\Delta}) \cap \mathbb{Z}^n)$, which contains every $\hat{\Delta}$ -mediated set.*

If $A(\hat{\Delta}) = \Delta^*$ resp. $\Delta^* = (\text{conv}(\hat{\Delta}) \cap \mathbb{Z}^n)$, we say that $\text{conv}(\hat{\Delta})$ is an M -simplex resp. H -simplex. It is not clear how to compute Δ^* efficiently and, hence, how to efficiently check whether a given simplex is an H -simplex. However, a brute force algorithm is given in [15].

Example 2.6. *The standard (Hurwitz) simplex given by $\text{conv}\{0, 2d \cdot e_1, \dots, 2d \cdot e_n\} \subset \mathbb{R}^n$ for $d \in \mathbb{N}$ is an H -simplex. The Newton polytope $\text{conv}\{0, (2,4), (4,2)\} \subset \mathbb{R}^2$ of the Motzkin polynomial $f = 1 + x^4y^2 + x^2y^4 - 3x^2y^2$ is an M -simplex (see Figure 1).*

In [10], we showed (generalizing results in [5, 15]) that the maximal $\hat{\Delta}$ -mediated sets Δ^* suffice to characterize the sums of squares property of polynomials as in Theorem 2.1.

Theorem 2.7 ([10]). *Let $f = \lambda_0 + \sum_{j=1}^n b_j x^{\alpha(j)} + c \cdot x^y \in \mathbb{R}[x]_{2d}$ be nonnegative and as in Theorem 2.1. For $c > 0$ and $y \in (2\mathbb{N})^n$, f is obviously a sum of (monomial) squares. In all other cases, the following equivalence does hold.*

$$f \text{ is a sum of squares} \Leftrightarrow f \text{ is a sum of binomial squares} \Leftrightarrow y \in \Delta^*.$$

Thus, for polynomials with H -simplex Newton polytope, nonnegativity and sums of squares coincide. Of course, with Theorem 2.3 this statement can be generalized immediately.

Corollary 2.8. *Let f be as in Theorem 2.3 such that $I \subseteq \Delta^*$. Then f is nonnegative if and only if f is a sum of binomial squares.*

Note that the conditions $y \in \Delta^*$ resp. $I \subseteq \Delta^*$ are in particular always satisfied if $\text{New}(f)$ is an H -simplex. Specifically, as pointed out in Example 2.6, the scaled standard (Hurwitz) simplex $2d\Delta_{n-1}$ used in [6] to propose a geometric program to produce lower bounds for polynomials is an H -simplex ([15]). Therefore, considering arbitrary H -simplices yields a significant extension of [6] and contains the scaled standard simplex $2d\Delta_{n-1}$ as one special instance. And considering polynomials with arbitrary simplex Newton polytope is an even further generalization.

In [10, Theorem 5.9.], sufficient conditions based on 2-normality of polytopes and toric geometry are given for a simplex to be an H -simplex. In particular, every “sufficiently large” simplex with even vertices is an H -simplex in \mathbb{R}^n (see [10, Section 5.1] for details).

3. MAIN RESULTS

In this section, we provide sufficient criteria on the coefficients of a polynomial to be a sum of nonnegative circuit polynomials as resp. a sum of squares and develop geometric programs capturing a richer class than in [6]. For the remainder of this article we make the following assumption.

Assumption: Let $f \in \mathbb{R}[x]_{2d}$ be a polynomial such that its Newton polytope $\text{New}(f)$ with vertices $\{0, \alpha(1), \dots, \alpha(n)\} \subset (2\mathbb{N})^n$ is a simplex.

Note that every $\alpha \in \Omega(f)$ can be written as a unique convex combination of the vertex set $\{0, \alpha(1), \dots, \alpha(n)\}$:

$$(3.1) \quad \alpha = \sum_{i=0}^n \lambda_i^{(\alpha)} \alpha(i) \quad \text{with} \quad \sum_{i=0}^n \lambda_i^{(\alpha)} = 1 \quad \text{and} \quad \lambda_i^{(\alpha)} \geq 0$$

where $\lambda_0^{(\alpha)}, \dots, \lambda_n^{(\alpha)} \in \mathbb{R}_{\geq 0}$ denote the scalars in the convex combination of $\alpha \in \Omega(f)$ in terms of the vertices of $\text{New}(f)$. Thus, we can write the polynomial f as

$$(3.2) \quad f = \sum_{\alpha \in \Omega(f)} \lambda_0^{(\alpha)} + \sum_{j=1}^n f_{\alpha(j)} x^{\alpha(j)} + \sum_{\alpha \in \Omega(f)} f_{\alpha} x^{\alpha}.$$

with $f_{\alpha(j)} > 0$ for $1 \leq j \leq n$ and $f_\alpha \in \mathbb{R}$. Scaling the polynomial by a new constant positive term $f_0 = \sum_{\alpha \in \Omega(f)} \lambda_0^{(\alpha)}$ is obviously irrelevant for nonnegativity of f and for polynomial optimization. The chosen scaling will turn out to be very suitable for our statements. In order to further simplify connections to the results in [5, 6] we consider the homogenized polynomial

$$F = \sum_{j=0}^n f_{\alpha(j)} x^{\alpha(j)} x_0^{2d-|\alpha(j)|} + \sum_{\alpha \in \Omega(f)} f_\alpha x^\alpha x_0^{2d-|\alpha|}$$

with $\alpha(0) = 0 \in \mathbb{N}^n$, $f_{\alpha(0)} = \sum_{\alpha \in \Omega(f)} \lambda_0^{(\alpha)}$ and $|\alpha| = \sum_{j=1}^n |\alpha_j| \in [0, 2d] \cap \mathbb{N}$ for all $\alpha \in \text{New}(f) \cap \mathbb{Z}^n$.

Theorem 3.1. *Let F be a homogeneous polynomial as above and suppose there exist $a_{\alpha,j} \geq 0$ for all $\alpha \in \Delta(F)$ and $0 \leq j \leq n$ such that*

- (1) $|f_\alpha| = \prod_{j=0}^n \left(\frac{a_{\alpha,j}}{\lambda_j^{(\alpha)}} \right)^{\lambda_j^{(\alpha)}}$ for all $\alpha \in \Delta(F)$.
- (2) $f_{\alpha(j)} \geq \sum_{\alpha \in \Delta(F)} a_{\alpha,j}$ for all $0 \leq j \leq n$.

Then $F - \sum_{\alpha \in \Omega(f) \setminus \Delta(f)} f_\alpha x^\alpha \cdot x_0^{2d-|\alpha|}$ and hence also F is a SONC. If, additionally, $\Delta(F) \subseteq \Delta^$, then F is a sum of binomial squares.*

Proof. Using Theorem 2.1, 2.7 and (1) we conclude that

$$\sum_{j=0}^n a_{\alpha,j} x^{\alpha(j)} x_0^{2d-|\alpha(j)|} + f_\alpha x^\alpha x_0^{2d-|\alpha|}$$

is a SONC for every $\alpha \in \Delta(F)$. By summing over all $\alpha \in \Delta(F)$ we have with Theorem 2.1 and Corollary 2.2 that

$$(3.3) \quad \sum_{j=0}^n \left(\sum_{\alpha \in \Delta(F)} a_{\alpha,j} \right) x^{\alpha(j)} x_0^{2d-|\alpha(j)|} + \sum_{\alpha \in \Delta(F)} f_\alpha x^\alpha x_0^{2d-|\alpha|}$$

is a SONC. Then condition (2) yields that

$$\sum_{j=0}^n f_{\alpha(j)} x^{\alpha(j)} x_0^{2d-|\alpha(j)|} + \sum_{\alpha \in \Delta(F)} f_\alpha x^\alpha x_0^{2d-|\alpha|}$$

is a SONC. Since for every $\alpha \in \Omega(F) \setminus \Delta(F)$ the term $f_\alpha x^\alpha$ is a monomial square, i.e., a nonnegative polynomial with a 0-simplex Newton polytope, F is a SONC. The sum of binomial square property follows from Theorem 2.7. \square

Setting $\alpha(j) = 2d$ for $0 \leq j \leq n$ the sum of binomial squares statement recovers [6, Theorem 2.3]. Additionally, e.g., by [6, Remark 2.4], we can assume that $a_{\alpha,j} = 0$ if and only if $\lambda_j^{(\alpha)} = 0$. Theorem 3.1 yields a new sufficient criterion on the coefficients of a polynomial to imply the SONC property as well as the sum of (binomial) squares property and significantly extends previous sums of squares criteria given in [5, 6, 11]. This extension

relies on the fact that the cited results assume the Newton polytope of the polynomial being the scaled standard simplex with degree $2d$, whereas Theorem 3.1 is, in particular, valid for *all* H -simplices containing the scaled standard simplex as a special instance.

Now, we describe the application of Theorem 3.1 in global optimization. To provide new lower bounds for polynomials, we prove our second main result.

Theorem 3.2. *Let $f \in \mathbb{R}[x]$ be a polynomial of degree $2d$ of the Form (3.2) with constant term $f_0 = \sum_{\alpha \in \Omega(f)} \lambda_0^{(\alpha)} > 0$ and let $r \in \mathbb{R}$. Suppose that for every $\alpha \in \Delta(f)$ there exist $a_{\alpha,1}, \dots, a_{\alpha,n} \geq 0$ with $a_{\alpha,j} = 0$ if and only if $\lambda_j^{(\alpha)} = 0$ (with $\lambda_j^{(\alpha)}$ in the sense of (3.1)) such that the following conditions hold.*

- (1) $|f_\alpha| = \prod_{j=1}^n \left(\frac{a_{\alpha,j}}{\lambda_j^{(\alpha)}} \right)^{\lambda_j^{(\alpha)}}$ for all $\alpha \in \Delta(f)$ with $|\alpha| = 2d$,
- (2) $f_{\alpha(j)} \geq \sum_{\alpha \in \Delta(f)} a_{\alpha,j}$ for all $1 \leq j \leq n$,
- (3) $f_0 - r \geq \sum_{\alpha \in \Delta^{<2d}(f)} \lambda_0^{(\alpha)} \cdot |f_\alpha|^{\frac{1}{\lambda_0^{(\alpha)}}} \cdot \prod_{j=1}^n \left(\frac{\lambda_j^{(\alpha)}}{a_{\alpha,j}} \right)^{\frac{\lambda_j^{(\alpha)}}{\lambda_0^{(\alpha)}}}$,

where $\Delta^{<2d}(f) = \{\alpha \in \Delta(f) : |\alpha| < 2d\}$. Then $f - r - \sum_{\alpha \in \Omega(f) \setminus \Delta(f)} f_\alpha x^\alpha$ and hence also $f - r$ is a SONC. If, additionally, $\Delta(f) \subseteq \Delta^*$, then $f - r$ is a sum of binomial squares.

Proof. We apply Theorem 3.1 to the homogenization of the polynomial $f - r$, which is given by

$$\overline{f - r} = (f_0 - r)x_0^{2d} + \sum_{j=1}^n f_{\alpha(j)} x^{\alpha(j)} x_0^{2d-|\alpha(j)|} + \sum_{\alpha \in \Delta(f)} f_\alpha x^\alpha x_0^{2d-|\alpha|}.$$

Then $\overline{f - r}$ is a SONC resp. a sum of binomial squares if only if $f - r$ is a SONC resp. a sum of binomial squares (see [6]). Our sufficient conditions in Theorem 3.1 now read as following.

- (1') $|f_\alpha| = \prod_{j=0}^n \left(\frac{a_{\alpha,j}}{\lambda_j^{(\alpha)}} \right)^{\lambda_j^{(\alpha)}} = \left(\frac{a_{\alpha,0}}{\lambda_0^{(\alpha)}} \right)^{\lambda_0^{(\alpha)}} \prod_{j=1}^n \left(\frac{a_{\alpha,j}}{\lambda_j^{(\alpha)}} \right)^{\lambda_j^{(\alpha)}}$,
- (2') $f_{\alpha(j)} \geq \sum_{\alpha \in \Delta(f)} a_{\alpha,j}$ for all $1 \leq j \leq n$ and $f_0 - r \geq \sum_{\alpha \in \Delta(f)} a_{\alpha,0}$.

Solving (1') for $a_{\alpha,0}$ yields

$$a_{\alpha,0} = \lambda_0^{(\alpha)} \cdot |f_\alpha|^{\frac{1}{\lambda_0^{(\alpha)}}} \cdot \prod_{j=1}^n \left(\frac{\lambda_j^{(\alpha)}}{a_{\alpha,j}} \right)^{\frac{\lambda_j^{(\alpha)}}{\lambda_0^{(\alpha)}}}$$

if $|\alpha| < 2d$. Set $a_{\alpha,0} = 0$ for $|\alpha| = 2d$. Conversely, defining $a_{\alpha,0}$ in this way, for every $\alpha \in \Delta$, one can verify that conditions (1) – (3) imply conditions (1'), (2') as follows: (2') follows immediately from (2) and (3) and definition of $a_{\alpha,0}$. Condition (1') follows from (1) again by homogenization and using definition of $a_{\alpha,0}$. \square

Again, setting $\alpha(j) = 2d$ for $1 \leq j \leq n$ the sum of binomial squares statement recovers [6, Theorem 3.1]. Now, we can define

$$f_{gp} = \sup \left\{ r \in \mathbb{R} : \begin{array}{l} \forall \alpha \in \Delta(f) \forall j \in \{1, \dots, n\} \exists a_{\alpha,j} \geq 0 \text{ with } a_{\alpha,j} = 0 \Leftrightarrow \lambda_j^{(\alpha)} = 0 \\ \text{such that conditions (1)-(3) of Theorem 3.2 are satisfied} \end{array} \right\}.$$

Indeed, f_{gp} is naturally connected to SONC certificates of nonnegativity as the following theorem shows.

Theorem 3.3. *Let $f \in \mathbb{R}[x]_{2d}$ be of the Form (3.2). Then*

$$f_{gp} = \sup \left\{ r \in \mathbb{R} : \begin{array}{l} \exists g_1, \dots, g_s \in \mathbb{R}[x]_{2d} \text{ with } \text{New}(f) = \text{New}(g_j) \text{ for } 1 \leq j \leq s \\ \text{and } f - r - \sum_{\alpha \in \Omega(f) \setminus \Delta(f)} f_{\alpha} x^{\alpha} = \sum_{j=1}^s g_j \text{ is a SONC} \end{array} \right\}.$$

Note in this context again that $\sum_{\alpha \in \Omega(f) \setminus \Delta(f)} f_{\alpha} x^{\alpha}$ is a sum of monomial squares, which is irrelevant for the computation of f_{gp} by Theorem 3.2.

Proof. By definition of f_{gp} and by Theorem 3.2, we already know that for every $r \geq f_{gp}$ it holds that $f - r - \sum_{\alpha \in \Omega(f) \setminus \Delta(f)} f_{\alpha} x^{\alpha}$ is a SONC and, by the Construction (3.3) in the proof of Theorem 3.1, we know that every polynomial g_j in the SONC decomposition satisfies $\text{New}(g_j) = \text{New}(f)$.

Hence, assume that there exist nonnegative circuit polynomials $g_1, \dots, g_s \in \mathbb{R}[x]_{2d}$ with $\text{New}(g_j) = \text{New}(f)$ for every j satisfying $f = \sum_{j=1}^s g_j$. W.l.o.g., we can assume that every $\alpha \in \Delta(f) \subset \text{New}(f)$ is contained in the support of a unique g_j – otherwise we can replace some $g_i + g_j$ by a new circuit polynomial g'_j . By Theorem 2.1, every g_j satisfies

$$g_j = \lambda_0^{\alpha(g_j)} + \sum_{i=1}^n g_{j,i} x^{\alpha(i)} + c_j x^{\alpha(g_j)}$$

with $\lambda_0^{\alpha(g_j)} \in \mathbb{R}_{\geq 0}, g_{i,j} \in \mathbb{R}_{>0}$ for all $1 \leq i \leq n$, $\alpha(g_j) \in \Delta(f) \subset \text{New}(f) \cap \mathbb{Z}^n$ and $|c_j| \leq \prod_{i=1}^n (g_{j,i} / \lambda_i^{(\alpha(g_j))}) \lambda_i^{(\alpha(g_j))}$. Hence, we have

$$f - r - \sum_{\alpha \in \Omega(f) \setminus \Delta(f)} f_{\alpha} x^{\alpha} = \sum_{j=1}^s g_j = \sum_{j=1}^s \lambda_0^{(\alpha(g_j))} + \sum_{i=1}^n \left(\sum_{j=1}^s g_{j,i} \right) x^{\alpha(i)} + \sum_{j=1}^s c_j x^{\alpha(g_j)}$$

satisfying conditions (1') and (2') in the proof of Theorem 3.2 and hence also conditions (1) – (3) of Theorem 3.2. \square

In the beginning of this article, we denoted that the observation of Ghasemi and Marshall was a trade off between fast solvability of the corresponding geometric programs in comparison with semidefinite programs and the fact that $f_{gp} \leq f_{sos}$. Here, we conclude the surprising fact that geometric programs do not have this lack in case of polynomials with simplex Newton polytope satisfying the conditions of Theorem 2.3. Quite the contrary, the bound f_{gp} will be at least as good as the bound f_{sos} . Note that the special instance $\#\Omega(f) = 1$ and $\text{New}(f)$ being the standard (Hurwitz-)simplex with edge length $2d$ was already observed by Ghasemi and Marshall (see [6, Corollary 3.4]).

Corollary 3.4. *Let f be a polynomial of the Form (3.2) with simplex Newton polytope $\text{New}(f) = \text{conv}\{0, \alpha(1), \dots, \alpha(n)\}$ and all $\alpha(j) \in (2\mathbb{N})^n$ such that $\Omega(f) \subseteq (\text{int}(\Delta) \cap \mathbb{N}^n)$ and there exists a vector $v \in (\mathbb{R}^*)^n$ such that $f_\alpha v^\alpha < 0$ for all $\alpha \in \Omega(f)$. Then*

$$f_{\text{sos}} \leq f_{\text{gp}} = f^*.$$

If additionally $\Delta(f) \subseteq \Delta^$, then*

$$f_{\text{sos}} = f_{\text{gp}}.$$

Particularly, if $\#\Omega(f) = 1$, then $f_{\text{gp}} = f_{\text{sos}} = f^$.*

Proof. The statement follows immediately from the Theorems 2.3 and 3.3 as well as Corollary 2.8. \square

4. GEOMETRIC PROGRAMMING

In this section, we prove that the number f_{gp} can indeed be obtained by a geometric program, which we introduce first.

Definition 4.1. *A function $f : \mathbb{R}_{>0}^n \rightarrow \mathbb{R}$ of the form $f(x) = f(x_1, \dots, x_n) = cx_1^{\alpha_1} \cdots x_n^{\alpha_n}$ with $c > 0$ and $\alpha_i \in \mathbb{R}$ is called a monomial (function). A sum $\sum_{i=0}^k c_i x_1^{\alpha_1(i)} \cdots x_n^{\alpha_n(i)}$ of monomials with $c_i > 0$ is called a posynomial (function).*

A geometric program has the following form.

$$(4.1) \quad \inf\{f_0(x) : x \in \mathbb{R}^n\} \quad \text{such that} \\ f_i(x) \leq 1 \quad \text{for all } 1 \leq i \leq m \quad \text{and} \quad g_j(x) = 1 \quad \text{for all } 1 \leq j \leq p,$$

where f_0, \dots, f_m are posynomials and g_1, \dots, g_p are monomial functions.

Geometric programs can be solved with an interior point method. In [14], the authors prove worst-case polynomial time complexity of this method. For an introduction and practical ability of geometric programs, see, e.g., [3, 4]. Based on our main results Theorem 3.1 and 3.2, we can draw the following corollary.

Corollary 4.2. *Let $f \in \mathbb{R}[x]$ be a non-constant polynomial of degree $2d$, $f_0 = \sum_{\alpha \in \Omega(f)} \lambda_0^{(\alpha)}$ with $f_{\alpha(j)} > 0$ for $1 \leq j \leq n$. Then $f_{\text{gp}} = f_0 - m^*$ where m^* is given by the following geometric program.*

$$\inf \left\{ \sum_{\alpha \in \Delta^{<2d}(f)} \lambda_0^{(\alpha)} \cdot |f_\alpha|^{\frac{1}{\lambda_0^{(\alpha)}}} \cdot \prod_{j=1}^n \left(\frac{\lambda_j^{(\alpha)}}{a_{\alpha,j}} \right)^{\frac{\lambda_j^{(\alpha)}}{\lambda_0^{(\alpha)}}} : \begin{array}{l} (a_{\alpha,1}, \dots, a_{\alpha,n}) \in \mathbb{R}_{\geq 0}^n \\ \text{for all } \alpha \in \Delta^{<2d}(f) \end{array} \right\} \quad \text{s.t.} \\ \sum_{\alpha \in \Delta(f)} \left(\frac{a_{\alpha,j}}{f_{\alpha(j)}} \right) \leq 1 \quad \text{for all } 1 \leq j \leq n \quad \text{and} \\ 1/|f_\alpha| \cdot \prod_{j=1}^n \left(\frac{a_{\alpha,j}}{\lambda_j^{(\alpha)}} \right)^{\lambda_j^{(\alpha)}} = 1 \quad \text{for all } \alpha \in \Delta(f) \quad \text{with } |\alpha| = 2d.$$

Proof. We have $f_{gp} = f_0 - m^*$ by definition of f_{gp} . Since

$$\sum_{\alpha \in \Delta^{<2d}(f)} \lambda_0^{(\alpha)} \cdot |f_\alpha|^{\frac{1}{\lambda_0^{(\alpha)}}} \cdot \prod_{j=1}^n \left(\frac{\lambda_j^{(\alpha)}}{a_{\alpha,j}} \right)^{\frac{\lambda_j^{(\alpha)}}{\lambda_0^{(\alpha)}}} \quad \text{and} \quad \sum_{\alpha \in \Delta(f)} \left(\frac{a_{\alpha,j}}{f_{\alpha(j)}} \right) \quad \text{for all } 1 \leq j \leq n$$

are posynomials in the variables $a_{\alpha,j}$ and for all $\alpha \in \Delta(f)$ with $|\alpha| = 2d$

$$1/|f_\alpha| \cdot \prod_{j=1}^n \left(\frac{a_{\alpha,j}}{\lambda_j^{(\alpha)}} \right)^{\lambda_j^{(\alpha)}}$$

is a monomial in the variables $a_{\alpha,j}$, m^* is indeed the output of a geometric program. \square

Corollary 4.3. *Let $\{(a_{\alpha,1}, \dots, a_{\alpha,n}) : \alpha \in \Delta(f)^{<2d}\}$ be the global minimizer of a geometric program as in Corollary 4.2. If $\Delta(f) = \Delta(f)^{<2d}$, then we have $\sum_{\alpha \in \Delta(f)} a_{\alpha,j} = f_{\alpha(j)}$ for every $1 \leq j \leq n$.*

Proof. Follows directly from Theorem 3.3 and Corollary 4.2. \square

4.1. Examples. We demonstrate our method and reflect our results by three examples. All following geometric programs are solved via the MATLAB solver GPPOSY¹.

- (1) First, consider the polynomial $f = \frac{1}{4} + x^8 + x^2y^6 + 4x^3y^3$. The geometric program proposed in [6] yields $f_{gp} = -\infty$, since the pure power y^8 is missing in the polynomial to make the Newton polytope a standard simplex of edge length 8. However, $\text{New}(f)$ is an H -simplex and we can use our results to compute f_{gp} . Here, we have $\Delta = \{\alpha\} = \{(3,3)\}$. Hence, we introduce the variables $a_{\alpha,j}$ for $j \in \{1,2\}$. Therefore, by Corollary 4.2, we have to solve the following geometric program.

$$\inf \left\{ \frac{1}{4} \cdot 4^4 \cdot \left(\frac{1}{4} \right)^{\frac{4}{4}} \cdot \left(\frac{1}{2} \right)^{\frac{4}{2}} \cdot a_{\alpha,1}^{-1} a_{\alpha,2}^{-2} : a_{\alpha,1} \leq 1, a_{\alpha,2} \leq 1 \right\}.$$

The optimal solution is given by $a_{\alpha,1} = a_{\alpha,2} = 1$ (as expected due to Corollary 4.3) yielding $m^* = 4$ and hence $f_{gp} = \frac{1}{4} - 4 = -3.75 = f_{sos} = f^*$ by Corollary 3.4.

- (2) Let $f = \frac{187}{208} + x^{80} + y^{78} - 8x^5y^3$. Again, the geometric program proposed in [6] yields $f_{gp} = -\infty$. But $\text{New}(f)$ is an H -simplex and with $\lambda_1^{(5,3)} = 1/16$ and $\lambda_2^{(5,3)} = 1/26$ our corresponding geometric program is given by

$$\inf \left\{ \frac{187}{208} \cdot 8^{\frac{208}{187}} \cdot \left(\frac{1}{16} \right)^{\frac{13}{187}} \cdot \left(\frac{1}{26} \right)^{\frac{8}{187}} \cdot a_{\alpha,1}^{-\frac{13}{187}} \cdot a_{\alpha,2}^{-\frac{8}{187}} : a_{\alpha,1} \leq 1, a_{\alpha,2} \leq 1 \right\}.$$

Using the software GLOPTIPOLY (see [8]), $f^* \approx -5.6179$ was computed in 4327, 2 seconds, i.e., approximately **1.2 hours**. However, using the above geometric program, we get a global minimizer $a_{\alpha,1} = a_{\alpha,2} = 1$ (again, as we would expect due

¹The Matlab version used was R2011a, running on a desktop computer with Intel(R) Core(TM)2 @ 2.33 GHz and 2 GB of RAM.

to Corollary 4.3) and the optimal solution $m^* = \frac{187}{208} \cdot \left(\frac{8^{208}}{16^{13} \cdot 26^8} \right)^{\frac{1}{187}}$ and, hence,
 $f^* = \lambda_0 - m^* = \frac{187}{208} \cdot \left(1 - \left(\frac{8^{208}}{16^{13} \cdot 26^8} \right)^{\frac{1}{187}} \right) \approx -5.6179$ in **0.5 seconds**.

- (3) Let now $f = \frac{17}{20} + 3x^8y^4 + 2x^6y^8 - 10x^3y^3 + x^5y^4$. Again, the geometric program in [6] cannot be used but the geometric program in Corollary 4.2 with $\Delta = \{\bar{\alpha}, \alpha\} = \{(3, 3), (5, 4)\}$ now reads as following.

$$\inf \left\{ \frac{9}{1250} \cdot 2^{\frac{1}{3}} \cdot 5^{\frac{2}{3}} \cdot a_{\alpha,1}^{-\frac{4}{3}} \cdot a_{\alpha,2}^{-1} + \frac{11}{40} \cdot 10^{\frac{3}{11}} \cdot 3^{\frac{9}{11}} \cdot 20^{\frac{8}{11}} (a_{\bar{\alpha},1})^{-\frac{3}{11}} (a_{\bar{\alpha},2})^{-\frac{6}{11}} \right\}$$

such that $\frac{a_{\alpha,1} + a_{\bar{\alpha},1}}{3} \leq 1$ and $\frac{a_{\alpha,2} + a_{\bar{\alpha},2}}{2} \leq 1$

Here, the variables $a_{\alpha,j}$ come from $\alpha = (5, 4)$ and $a_{\bar{\alpha},j}$ come from $\bar{\alpha} = (3, 3)$. Again, we use the MATLAB solver GPPOSY to solve this geometric program with the following code:

```
>> A0=[-4/3,-1,0,0;0,0,-3/11,-6/11]
>> A1=[1,0,0,0;0,0,1,0]
>> A2=[0,1,0,0;0,0,0,1]
>> A=[A0;A1;A2]
>> b0=[9/1250*2^(1/3)*5^(2/3);11/40*10^(3/11)*3^(9/11)*20^(8/11)]
>> b1=[1/3;1/3]
>> b2=[1/2;1/2]
>> b=[b0;b1;b2]
>> szs=[size(A0,1);size(A1,1);size(A2,1)]
>> [x,status,lambda,nu]=gpposy(A,b,szs)
```

The optimal solution is given by

$$(a_{\alpha,1}, a_{\alpha,2}, a_{\bar{\alpha},1}, a_{\bar{\alpha},2}) = (0.5910, 0.1685, 2.4090, 1.8315)$$

(Corollary 4.3 holds again) yielding $m^* \approx -6.644$ and hence $f_{gp} = \frac{17}{20} - 6.644 \approx -5.794$. Interestingly, we also have $f_{sos} = f^* \approx -5.794$ by using GLOPTIPOLY. Note that $f_{gp} = f^*$ by Theorem 2.3.

- (4) The Motzkin polynomial $f = \frac{1}{3} + \frac{1}{3}x^4y^2 + \frac{1}{3}x^2y^4 - x^2y^2$ satisfies $f_{gp} = f^* = 0$ again by Theorem 2.3. However, $f_{sos} = -\infty$.

- (5) Let $f = \frac{5}{12} + \frac{5}{24}x^6 + \frac{5}{24}x^2y^4 + \frac{5}{24}x^2y^2 - \frac{5}{8}xy$. Then one can check that
 $f_{gp} \approx -0.41 < f_{sos} = f^* \approx 0.196$.

5. CONCLUSION AND OUTLOOK

We have proposed a new geometric program for producing lower bounds for polynomials that extends the existing one in [6]. This extension sheds light to the crucial structure of the Newton polytope of polynomials. In particular, our results serve as a next step in optimization of polynomials with simplex Newton polytopes and connect this problem to

- (1) the recently established SONC nonnegativity certificates, and
- (2) the construction of simplices with an interesting lattice point structure, namely, what we have called H -simplices in this article.

The trade off in [6] between the bound f_{gp} being worse than f_{sos} but f_{gp} allowing to solve higher dimensional examples in rather quick time cannot be observed in our refinement. Since fast solvability of the geometric programs still holds, the bounds f_{gp} and f_{sos} are not comparable. There are classes for which $f_{gp} \leq f_{sos}$ holds ([6]), but also classes with $f_{gp} \geq f_{sos}$ (Theorem 2.3), which, due to fast solvability, is a very delicate case. It would be interesting to classify more classes for which the bounds are comparable. Hence, an analysis of the gap $f_{sos} - f_{gp}$ is an interesting task having major impact on computational complexity of solving polynomial optimization problems. Equivalently, looking from a convex geometric viewpoint, it would be interesting to analyze the gap between the cone of sums of squares and the cone of sums of binomial squares as well as the gap between the cone of sums of squares and the cone of sums of nonnegative circuit polynomials.

ACKNOWLEDGMENTS

The second author was supported by DFG grant MA 4797/3-2.

REFERENCES

1. G. Blekherman, P.A. Parrilo, and R.R. Thomas, *Semidefinite optimization and convex algebraic geometry*, MOS-SIAM Series on Optimization, vol. 13, SIAM and the Mathematical Optimization Society, Philadelphia, 2013.
2. L. Blum, F. Cucker, M. Shub, and S. Smale, *Complexity and real computation*, Springer-Verlag, New York, 1998.
3. S. Boyd, S.J. Kim, L. Vandenberghe, and A. Hassibi, *A tutorial on geometric programming*, Optim. Eng. **8** (2007), no. 1, 67–127.
4. S. Boyd and L. Vandenberghe, *Convex optimization*, Cambridge University Press, Cambridge, 2004.
5. C. Fidalgo and A. Kovacec, *Positive semidefinite diagonal minus tail forms are sums of squares*, Math. Z. **269** (2011), no. 3-4, 629–645.
6. M. Ghasemi and M. Marshall, *Lower bounds for polynomials using geometric programming*, SIAM J. Optim. **22** (2012), no. 2, 460–473.
7. ———, *Lower bounds for a polynomial on a basic closed semialgebraic set using geometric programming*, (2013), Preprint, arxiv:1311.3726.
8. D. Henrion, J.B. Lasserre, and J. Löfberg, *GloptiPoly 3: moments, optimization and semidefinite programming*, Optim. Methods Softw. **24** (2009), no. 4-5, 761–779.
9. D. Hilbert, *Ueber die Darstellung definiter Formen als Summe von Formenquadraten*, Math. Ann. **32** (1888), no. 3, 342–350.
10. S. Ilman and T. de Wolff, *Amoebas, nonnegative polynomials and sums of squares supported on circuits*, 2014, Preprint, arXiv:1402.0462.
11. J.B. Lasserre, *Sufficient conditions for a real polynomial to be a sum of squares*, Arch. Math. (Basel) **89** (2007), no. 5, 390–398.
12. ———, *Moments, positive polynomials and their applications*, Imperial College Press Optimization Series, vol. 1, Imperial College Press, London, 2010.
13. M. Laurent, *Sums of squares, moment matrices and optimization over polynomials*, Emerging applications of algebraic geometry, IMA Vol. Math. Appl., vol. 149, Springer, New York, 2009, pp. 157–270.
14. Y. Nesterov and A. Nemirovskii, *Interior point polynomial algorithms in convex programming*, Studies in Applied Mathematics, Society for Industrial and Applied Mathematics, 1994.

15. B. Reznick, *Forms derived from the arithmetic-geometric inequality*, Math. Ann. **283** (1989), no. 3, 431–464.

SADIK ILIMAN, GOETHE-UNIVERSITÄT, FB 12 – INSTITUT FÜR MATHEMATIK, POSTFACH 11 19 32, 60054 FRANKFURT AM MAIN, GERMANY

TIMO DE WOLFF, UNIVERSITÄT DES SAARLANDES, FACHRICHTUNG MATHEMATIK, POSTFACH 15 11 50, 66041 SAARBRÜCKEN, GERMANY

E-mail address: `iliman@math.uni-frankfurt.de`, `dewolff@math.uni-sb.de`